

**Computer Engineering Department
Bilkent University**

**CS533 – Information Retrieval Systems
May 21, 2011**

FINAL EXAM STUDY GUIDE

A. Exam Structure and Time Schedule

1. Similar to the midterm exam.
2. No. of questions 7-8 (approx.) with subsections.
3. Final exam is on May 23, Monday, at 3:40-5:30 pm, place EA502.

B. Subjects to be Covered (an incomplete lists)

1. Midterm topics
2. TDV: term discrimination value
3. Term weighting for query processing (Salton, Buckley paper; studied in class in detail)
4. Stemming: n-gram based, successor variety
5. PAT String and PAT Arrays: construction and use
6. Signature Files: sequential, bit sliced, tree structured signature files, signature use in database applications, fixed and other related partitioning methods, LHSS (Linear Hashing with Superimposed Signatures, also known as Quick Filter)
7. Index/data structures for information filtering
8. Papers: does not require cover to cover reading; just know the fundamental concepts (~15%).

NOTES FOR PROJECTS

1. Change in project report submission deadline: You have an extension, May 25, 2011, Wednesday by noon time is the new deadline. Project report in the form of ACM conference papers (<http://www.acm.org/sigs/publications/proceedings-templates>, ~8 pages). It involves submission of the developed programs with a short report for the programs (their structure, etc.). No need to exaggerate, a short program report is enough. You may be required to do a demo.
2. You need to submit a hardcopy of your project and also its pdf form and also an updated form of your class presentation with your results. You need to do these submissions again by noon time on May 25.
3. On May 25 and 26 please find me in my office for a demo etc. of your project.